

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:

(ii) TITLE OF INVENTION: Novel CREBa Isoform

(iii) NUMBER OF SEQUENCES: 9

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: MArshall, O'Toole, Gerstein, Murray & Borun
(B) STREET: 233 South Wacker Drvie, 6300 Sears Tower
(C) CITY: Chicago
(D) STATE: Illinois
(E) COUNTRY: USA
(F) ZIP: 60606

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE: Januray 12, 1998
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Williams Jr., Joseph A.
(B) REGISTRATION NUMBER: 38,659
(C) REFERENCE/DOCKET NUMBER: 27866/33469

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: 312-474-6300
(B) TELEFAX: 312-474-0448

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 3190 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 304..1866

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GGCACGAGGG ACTTTCTTGG GATGAGCGCT GCCTTTTGG CTTCCCTTTG GATGCACAGC	60
CCGATTTAAC CCCTGCACCT TCCGCCGAT CCCAGCAGGC TTGTCCTCCC CGGGGAGTCA	120
CAGATTTCCG AGGACAAGGG TCGCGTAGCC TTCGGCAGGG CTCTCCCGAG TTCCTGCTCC	180
AGTGCATAAG TTCCACGCGC GCACACGCCA AGTACACGGG GAGAACGTC TCACCGGCC	240
GCGGCGGCTC TGCGCGGTCC CCTCCTGCCT CAGCATCCTC GGGCCTGCGC GGCGCCACC	300
GCC ATG GAG GTG CTG GAG AGC GGG GAG CAG AGC GTC CTG CAG TGG GAC	348
Met Glu Val Leu Glu Ser Gly Glu Gln Ser Val Leu Gln Trp Asp	
1 5 10 15	
CGC AAG CTG AGC GAG CTG TCA GAG CCC GGA GAG ACT GAG GCC CTC ATG	396
Arg Lys Leu Ser Glu Leu Ser Glu Pro Gly Glu Thr Glu Ala Leu Met	
20 25 30	
TAC CAC ACG CAC TTC TCG GAG CTC CTA GAC GAG TTT TCC CAG AAC GTC	444
Tyr His Thr His Phe Ser Glu Leu Leu Asp Glu Phe Ser Gln Asn Val	
35 40 45	
CTG GGT CAG CTC CTG AGT GAC CCT TTC CTC TCA GAG AAG AGC GAG TCA	492
Leu Gly Gln Leu Leu Ser Asp Pro Phe Leu Ser Glu Lys Ser Glu Ser	
50 55 60	
ATG GAG GTG GAG CCA TCT CCA ACA TCA CCA GCG CCT CTC ATC CAG GCT	540
Met Glu Val Glu Pro Ser Pro Thr Ser Pro Ala Pro Leu Ile Gln Ala	
65 70 75	
GAA CAC AGC TAC TCT CTG AGC GAG GAG CCC CGG ACT CAG TCA CCA TTT	588
Glu His Ser Tyr Ser Leu Ser Glu Glu Pro Arg Thr Gln Ser Pro Phe	
80 85 90 95	
ACC CAT GCG GCT ACC AGC GAC AGC TTC AAT GAC GAG GAG GTG GAG AGT	636
Thr His Ala Ala Thr Ser Asp Ser Phe Asn Asp Glu Glu Val Glu Ser	
100 105 110	
GAA AAA TGG TAC CTG TCT ACA GAG TTT CCT TCA GCT ACC ATC AAG AAA	684
Glu Lys Trp Tyr Leu Ser Thr Glu Phe Pro Ser Ala Thr Ile Lys Lys	
115 120 125	
GAG CCA ATC ACA GAG GAG CAG CCC CCG GGA CTT GTC CCT TCT GTC ACT	732
Glu Pro Ile Thr Glu Glu Gln Pro Pro Gly Leu Val Pro Ser Val Thr	
130 135 140	
CTG ACC ATC ACA GCC ATT TCC ACT CCT TTT GAA AAA GAA GAG TCC CCT	780
Leu Thr Ile Thr Ala Ile Ser Thr Pro Phe Glu Lys Glu Glu Ser Pro	
145 150 155	
CTG GAT ATG AAT GCT GGG GGG GAC TCC TCA TGC CAG ACG CTT ATT CCT	828

Leu Asp Met Asn Ala Gly Gly Asp Ser Ser Cys Gln Thr Leu Ile Pro
160 165 170 175

AAG ATT AAG CTG GAG CCC CAC GAA GTG GAT CAG TTC TTA AAC TTC TCC 876
Lys Ile Lys Leu Glu Pro His Glu Val Asp Gln Phe Leu Asn Phe Ser
180 185 190

CCG AAA GAA GCC TCC GTG GAT CAA CTG CAC TTA CCA CCA ACA CCA CCC 924
Pro Lys Glu Ala Ser Val Asp Gln Leu His Leu Pro Pro Thr Pro Pro
195 200 205

AGT AGT CAC AGC AGT GAC TCT GAG GGC AGC TTG AGC CCC AAC CCA CGC 972
Ser Ser His Ser Ser Asp Ser Glu Gly Ser Leu Ser Pro Asn Pro Arg
210 215 220

CTG CAT CCC TTC AGC CTG TCT CAG GCC CAC AGC CCT GTC AGA GCC ATG 1020
Leu His Pro Phe Ser Leu Ser Gln Ala His Ser Pro Val Arg Ala Met
225 230 235

CCC CGG GGC CCC TCT GCC TTG TCC ACA TCT CCT CTC CTC ACA GCT CCA 1068
Pro Arg Gly Pro Ser Ala Leu Ser Thr Ser Pro Leu Leu Thr Ala Pro
240 245 250 255

CAT AAG CTG CAG GGA TCG GGC CCC CTG GTC CTG ACA GAA GAG GAG AAG 1116
His Lys Leu Gln Gly Ser Gly Pro Leu Val Leu Thr Glu Glu Glu Lys
260 265 270

AGG ACC CTG GTT GCC GAG GGC TAT CCC ATT CCC ACC AAG CTG CCT CTG 1164
Arg Thr Leu Val Ala Glu Gly Tyr Pro Ile Pro Thr Lys Leu Pro Leu
275 280 285

ACA AAA TCT GAG GAG AAG GCC CTG AAG AAA ATC CGG AGA AAG ATC AAG 1212
Thr Lys Ser Glu Glu Lys Ala Leu Lys Ile Arg Arg Lys Ile Lys
290 295 300

AAT AAG ATT TCT GCC CAA GAA AGC AGG AGA AAG AAG AAA GAA TAC ATG 1260
Asn Lys Ile Ser Ala Gln Glu Ser Arg Arg Lys Lys Glu Tyr Met
305 310 315

GAC AGC CTG GAG AAA AAA GTG GAG TCT TGT TCA ACT GAG AAC TTG GAG 1308
Asp Ser Leu Glu Lys Lys Val Glu Ser Cys Ser Thr Glu Asn Leu Glu
320 325 330 335

CTT CGG AAG AAG GTG GAG GTG CTG GAG AAC ACC AAT AGG ACT CTC CTT 1356
Leu Arg Lys Lys Val Glu Val Leu Glu Asn Thr Asn Arg Thr Leu Leu
340 345 350

CAG CAA CTT CAG AAG CTT CAG ACT TTG GTG ATG GGG AAG GTC TCT CGA 1404
Gln Gln Leu Gln Lys Leu Gln Thr Leu Val Met Gly Lys Val Ser Arg
355 360 365

ACC TGC AAG TTA GCT GGC ACA CAG ACT GGC ACC TGC CTC ATG GTC GTT 1452
Thr Cys Lys Leu Ala Gly Thr Gln Thr Gly Thr Cys Leu Met Val Val
370 375 380

GTG CTT TGC TTT GCT GTT GCA TTT GGA AGC TTC TTT CAA GGC TAT GGG Val Leu Cys Phe Ala Val Ala Phe Gly Ser Phe Phe Gln Gly Tyr Gly 385 390 395	1500
CCT TAT CCT TCT GCC ACC AAG ATG GCT CTG CCC AGC CAG CAT CCT CTG Pro Tyr Pro Ser Ala Thr Lys Met Ala Leu Pro Ser Gln His Pro Leu 400 405 410 415	1548
TCA GAG CCA TAC ACA GCC TCC GTG GTG AGA TCC AGG AAC CTG CTA ATC Ser Glu Pro Tyr Thr Ala Ser Val Val Arg Ser Arg Asn Leu Leu Ile 420 425 430	1596
TAT GAG GAA CAC GCT CCC CTG GAA GAG TCG TCG AGC CCA GCC TCA ACC Tyr Glu His Ala Pro Leu Glu Ser Ser Ser Pro Ala Ser Thr 435 440 445	1644
GGG GAG CTG GGG GGC TGG GAC AGA GGC TCC TCT CTG CTC AGG GCA TCG Gly Glu Leu Gly Trp Asp Arg Gly Ser Ser Leu Leu Arg Ala Ser 450 455 460	1692
TCG GGG CTT GAG GCC CTG CCA GAG GTG GAT CTT CCC CAT TTC CTT ATC Ser Gly Leu Glu Ala Leu Pro Glu Val Asp Leu Pro His Phe Leu Ile 465 470 475	1740
TCC AAT GAG ACG AGC TTG GAG AAG TCA GTA CTG TTG GAG CTT CAG CAG Ser Asn Glu Thr Ser Leu Glu Lys Ser Val Leu Leu Glu Leu Gln Gln 480 485 490 495	1788
CAC CTG GTC AGC AGC AAA CTG GAA GGG AAC GAA ACA CTC AAG GTT GTA His Leu Val Ser Ser Lys Leu Glu Gly Asn Glu Thr Leu Lys Val Val 500 505 510	1836
GAG CTG GAG AGG AGA GTG AAC GCC ACC TTC TGAGGAGAGC TCCACCCCTCC Glu Leu Glu Arg Arg Val Asn Ala Thr Phe 515 520	1886
TCTTCTCCTA ACTCCATCTG ATCGTCCTTT CAGTTCCCC TTCACCACTG GATCTCGAGG AGGAGATGGC TAGTGTACG GCTCGAGACA GGAGGCCAGC CCAGGGGGTT CTGCTTATGT	1946
GTCCCCGTGG CTCTCCACAA AAGGGAGCTA GCACCTCTCC ATCCCTTTCT CTTACTGCCA	2006
TTGGAAATTA TTTTAGGGCT GAGATAGGGG TGGAACGAGC AGGCTTGTGTT CCACCAATAG	2126
TGCCAAGAAG ACACGTGCTG ATTCTTCCCC GGGAGGGAGTG ACTCCTCTGA AGAAGACATG	2186
ACTCATGTTG AGTTGAGACC CCAGACTCTA GCCACACACA TGCCACAGAC ATGCCAGGGAA	2246
GTGGCAAAGC ACTGACTCCT GAGCTCCCTT CCTCACTAGG ACTCCAGTGT GACCCTGCAC	2306
TGAGAGGACC AAAGCGTCAT TGCAGTCTTC TCTCCACCCCT GTACCCCGGA GTCCTGATTG	2366
GATGTCTGCA GAGGCAGATG GGGCTCCAC CATATTTCA GGCCGCAAGT GCAATTCCCTG	2426

AAGGCATCAG GCTCTCTCT CCCAGGCTCT CCTGCCACT GTGTTGTTG TAGGACACCC 2486
CCACACCCAC TCATACACAG CCTGCATCTC CACAGGACAA TAGCTCTGTC TCCCTGGCCT 2546
CCCCTCCCCA TTTGTAAATA GTATTTATTA GCTTGCTCAA GCTCCCAGCT GGCCATAGTG 2606
AAAAGATTTC CCCTTCAAC CAGCAAAGTC TTCTGTTGGC CTTTGGAAACA GGAGAGTCCC 2666
CGGAATCTAG GACCCTAGTC TTTGTACTTG ATGCCTTGT TCCCCCTTT TCTTCTTTAA 2726
AATTGGGGAC CTATAACATC ATCGCTGTT CGGAATCCAC TTAGGCATGT GTCCCCTGAT 2786
GGATGAATAC ATGGGAATGG TGGATACTGT CTTCTGACTC AGGCTCTAGG CTCCATGGCT 2846
TCCTCTCTCT GGTCTGCCA CACAGAAGGA AAGCCCTGTC CAGGATAATG AGCGTTGCTG 2906
ACACCCTTGC TAGCTTGTCC TGCCTACCTG CTTACCCAC TCCCTCACCT TCCTCCTTCC 2966
CTTCTGCCCT CCATCCACCT GCCTTAACTA ATTGGGGCTG GAGTTGGTCA TTTTTTGTAC 3026
ACCCACAGTG GTACCTTTA CAGTCAGGTT TGGATACTTT GCAGCTCATC CAAAGAGACA 3086
TAACTAAACC CTAAACTCTT TTTTGTTGT TGTTGTTGT GTTTTTTTT TTTATGATTA 3146
AAAAGTAAAA ATTGTAGTTT AAAAAAAAAA AAAAAAAACT CGAG 3190

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 521 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Glu Val Leu Glu Ser Gly Glu Gln Ser Val Leu Gln Trp Asp Arg
1 5 10 15

Lys Leu Ser Glu Leu Ser Glu Pro Gly Glu Thr Glu Ala Leu Met Tyr
20 25 30

His Thr His Phe Ser Glu Leu Leu Asp Glu Phe Ser Gln Asn Val Leu
35 40 45

Gly Gln Leu Leu Ser Asp Pro Phe Leu Ser Glu Lys Ser Glu Ser Met
50 55 60

Glu Val Glu Pro Ser Pro Thr Ser Pro Ala Pro Leu Ile Gln Ala Glu
65 70 75 80

His Ser Tyr Ser Leu Ser Glu Glu Pro Arg Thr Gln Ser Pro Phe Thr

85

90

95

His Ala Ala Thr Ser Asp Ser Phe Asn Asp Glu Glu Val Glu Ser Glu
100 105 110

Lys Trp Tyr Leu Ser Thr Glu Phe Pro Ser Ala Thr Ile Lys Lys Glu
115 120 125

Pro Ile Thr Glu Glu Gln Pro Pro Gly Leu Val Pro Ser Val Thr Leu
130 135 140

Thr Ile Thr Ala Ile Ser Thr Pro Phe Glu Lys Glu Glu Ser Pro Leu
145 150 155 160

Asp Met Asn Ala Gly Gly Asp Ser Ser Cys Gln Thr Leu Ile Pro Lys
165 170 175

Ile Lys Leu Glu Pro His Glu Val Asp Gln Phe Leu Asn Phe Ser Pro
180 185 190

Lys Glu Ala Ser Val Asp Gln Leu His Leu Pro Pro Thr Pro Pro Ser
195 200 205

Ser His Ser Ser Asp Ser Glu Gly Ser Leu Ser Pro Asn Pro Arg Leu
210 215 220

His Pro Phe Ser Leu Ser Gln Ala His Ser Pro Val Arg Ala Met Pro
225 230 235 240

Arg Gly Pro Ser Ala Leu Ser Thr Ser Pro Leu Leu Thr Ala Pro His
245 250 255

Lys Leu Gln Gly Ser Gly Pro Leu Val Leu Thr Glu Glu Glu Lys Arg
260 265 270

Thr Leu Val Ala Glu Gly Tyr Pro Ile Pro Thr Lys Leu Pro Leu Thr
275 280 285

Lys Ser Glu Glu Lys Ala Leu Lys Lys Ile Arg Arg Lys Ile Lys Asn
290 295 300

Lys Ile Ser Ala Gln Glu Ser Arg Arg Lys Lys Glu Tyr Met Asp
305 310 315 320

Ser Leu Glu Lys Lys Val Glu Ser Cys Ser Thr Glu Asn Leu Glu Leu
325 330 335

Arg Lys Lys Val Glu Val Leu Glu Asn Thr Asn Arg Thr Leu Leu Gln
340 345 350

Gln Leu Gln Lys Leu Gln Thr Leu Val Met Gly Lys Val Ser Arg Thr
355 360 365

Cys Lys Leu Ala Gly Thr Gln Thr Gly Thr Cys Leu Met Val Val Val

370 375 380
Leu Cys Phe Ala Val Ala Phe Gly Ser Phe Phe Gln Gly Tyr Gly Pro
385 390 395 400
Tyr Pro Ser Ala Thr Lys Met Ala Leu Pro Ser Gln His Pro Leu Ser
405 410 415
Glu Pro Tyr Thr Ala Ser Val Val Arg Ser Arg Asn Leu Leu Ile Tyr
420 425 430
Glu Glu His Ala Pro Leu Glu Glu Ser Ser Ser Pro Ala Ser Thr Gly
435 440 445
Glu Leu Gly Gly Trp Asp Arg Gly Ser Ser Leu Leu Arg Ala Ser Ser
450 455 460
Gly Leu Glu Ala Leu Pro Glu Val Asp Leu Pro His Phe Leu Ile Ser
465 470 475 480
Asn Glu Thr Ser Leu Glu Lys Ser Val Leu Leu Glu Leu Gln Gln His
485 490 495
Leu Val Ser Ser Lys Leu Glu Gly Asn Glu Thr Leu Lys Val Val Glu
500 505 510
Leu Glu Arg Arg Val Asn Ala Thr Phe
515 520

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CGCGGATCCT AATGGAGCTG AGAGTCGGG

29

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

CGCGGGATCCG CTCATCGGTG CACGACAGA

29

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 29 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGGGATCCTC ACAGACTCCAC ATAAGCTGC

29

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 28 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGAATTCGCT CAAGGAGAGT CCTATTGG

28

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 154 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CAGGTCAGTT CAGCGGATCC TGTGNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	60
NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN	120
NNNNNNNNNN AGGCGAATTC AGTGCAACTG CAGC	154

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CAGGTCAGTT CAGCGGATCC TGTGCG	25
------------------------------	----

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 25 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCTGCAGTTG CACTGAATTC GCCTC	25
-----------------------------	----

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 38 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

GAATCGGGCC GCCGAGATCT CATATGGAGC TGAGAGTC

38